

SEQUENCE LISTING

(1) GENERAL INFORMATION:

(i) APPLICANT: Band, Vimla

(ii) TITLE OF INVENTION: NES-1 POLYPEPTIDES, DNA, AND RELATED MOLECULES AND METHODS

(iii) NUMBER OF SEQUENCES: 11

(iv) CORRESPONDENCE ADDRESS:

- (A) ADDRESSEE: Fish & Richardson P.C.
- (B) STREET: 225 Franklin Street
- (C) CITY: Boston
- (D) STATE: MA
- (E) COUNTRY: USA
- (F) ZIP: 02110-2804

(v) COMPUTER READABLE FORM:

- (A) MEDIUM TYPE: Floppy disk
- (B) COMPUTER: IBM PC compatible
- (C) OPERATING SYSTEM: PC-DOS/MS-DOS
- (D) SOFTWARE: PatentIn Release #1.0, Version #1.30

(vi) CURRENT APPLICATION DATA:

- (A) APPLICATION NUMBER:
- (B) FILING DATE:

(vii) PRIOR APPLICATION DATA:

- (A) APPLICATION NUMBER: 08/467,155
- (B) FILING DATE: 06-JUN-1995
- (C) CLASSIFICATION:

(viii) ATTORNEY/AGENT INFORMATION:

- (A) NAME: Clark, Paul T.
- (B) REGISTRATION NUMBER: 30,162
- (C) REFERENCE/DOCKET NUMBER: 00398/100002

(ix) TELECOMMUNICATION INFORMATION:

- (A) TELEPHONE: 617/542-5070
- (B) TELEFAX: 617/542-8906
- (C) TELEX: 200154

(2) INFORMATION FOR SEQ ID NO:1:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 276 amino acids
- (B) TYPE: amino acid
- (C) STRANDEDNESS: not relevant
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:1:

Met Arg Ala Pro His Leu His Leu Ser Ala Ala Ser Gly Ala Arg Ala
1 5 10 15

Ala Ala Leu Leu Pro Gln Asn Asp Thr Arg Leu Asp Pro Glu Ala Tyr
 35 40 45
 Gly Ala Pro Cys Ala Arg Gly Ser Gln Pro Trp Gln Val Ser Leu Phe
 50 55 60
 Asn Gly Leu Ser Phe His Cys Ala Gly Val Leu Val Asp Gln Ser Trp
 65 70 75 80
 Val Leu Thr Ala Ala His Cys Gly Asn Lys Pro Leu Trp Ala Arg Val
 85 90 95
 Gly Asp Asp His Leu Leu Leu Gln Gly Glu Gln Leu Arg Arg Thr
 100 105 110
 Thr Arg Ser Val Val His Pro Lys Tyr His Gln Gly Ser Gly Pro Ile
 115 120 125
 Leu Pro Arg Arg Thr Asp Glu His Asp Leu Met Leu Leu Lys Leu Ala
 130 135 140
 Arg Pro Val Val Pro Gly Pro Arg Val Arg Ala Leu Gln Leu Pro Tyr
 145 150 155 160
 Arg Cys Ala Gln Pro Gly Asp Gln Cys Gln Val Ala Gly Trp Gly Thr
 165 170 175
 Thr Ala Ala Arg Arg Val Lys Tyr Asn Lys Gly Leu Thr Cys Ser Ser
 180 185 190
 Ile Thr Ile Leu Ser Pro Lys Glu Cys Glu Val Phe Tyr Pro Gly Val
 195 200 205
 Val Thr Asn Asn Met Ile Cys Ala Gly Leu Asp Arg Gly Gln Asp Pro
 210 215 220
 Cys Gln Ser Asp Ser Gly Gly Pro Leu Val Cys Asp Glu Thr Leu Gln
 225 230 235 240
 Gly Ile Leu Ser Trp Gly Val Tyr Pro Cys Gly Ser Ala Gln His Pro
 245 250 255
 Ala Val Tyr Thr Gln Ile Cys Lys Tyr Met Ser Trp Ile Asn Lys Val
 260 265 270
 Ile Arg Ser Asn
 275

(2) INFORMATION FOR SEQ ID NO:2:

- (i) SEQUENCE CHARACTERISTICS:
 - (A) LENGTH: 1454 base pairs
 - (B) TYPE: nucleic acid
 - (C) STRANDEDNESS: single
 - (D) TOPOLOGY: linear
- (ii) MOLECULE TYPE: cDNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:2:

ACCAGCGGCA GACCACAGGC AGGGCAGAGG CACGTCTGGG TCCCCCTCCCT CCTTCCTATC

SEQUENCE DATA

GGCGACTCCC	AGATCCTGGC	CATGAGAGCT	CCGCACCTCC	ACCTCTCCGC	CGCCTCTGGC	120
GCCCCGGCTC	TGGCGAAGCT	GCTGCCGCTG	CTGATGGCGC	AACTCTGGC	CGCAGAGGGC	180
GCGCTGCTCC	CCCAAAACGA	CACGCCGTTG	GACCCCGAAG	CCTATGGCGC	CCCGTGCACG	240
CGCGGCTCGC	AGCCCTGGCA	GGTCTCGCTC	TTCAACGGCC	TCTCGTTCCA	CTGCGGGGT	300
GTCTGGTGG	ACCAGAGTTG	GGTGTGACG	GCCGCCACT	GCGGAAACAA	GCCACTGTGG	360
GCTCGAGTAG	GGGATGATCA	CCTGCTGCTT	CTTCAGGGCG	ACCAGCTCCG	CCGGACGACT	420
CGCTCTGTTG	TCCATCCCAA	GTACCACCAAG	GGCTCAGGCC	CCATCCTGCC	AAGGCGAACG	480
GATGAGCACG	ATCTCATGTT	GCTAAAGCTG	GCCAGGCCCG	TAGTGCCGGG	GCCCCGCGTC	540
CGGGCCCTGC	AGCTTCCCTA	CCGCTGTGCT	CAGCCCCGAG	ACCAAGTGCCTA	GGTTGCTGGC	600
TGGGGCACCA	CGGCGCCCG	GAGAGTGAAG	TACAACAAGG	GCCTGACCTG	CTCCAGCCTC	660
ACTATCCTGA	GCCCTAAAGA	GTGTGAGGTC	TTCTACCTG	GCGTGGTCAC	CAACAACATG	720
ATATGTGCTG	GAATGGACCG	GGGCCAGGAC	CCTTGCCAGA	GTGACTCTGG	AGGCCCCCTG	780
GTCTGTGACG	AGACCCCTCA	AGGCATCCTC	TCGTGGGTG	TTTACCCCTG	TGGCTCTGCC	840
CAGCATCCAG	CTGTCTACAC	CCAGATCTGC	AAATACATGT	CCTGGATCAA	TAAAGTCATA	900
CGCTCCAAC	GATCCAGATG	CTACGCTCCA	GCTGATCCAG	ATGTTATGCT	CCTGCTGATC	960
CAGATGCCCA	GAGGCTCCAT	CGTCCATCCT	CTTCCTCCCC	AGTCGGCTGA	ACTCTCCCT	1020
TGTCTGCACT	GTTCAAAACCT	CTGCCGCCCT	CCACACCTCT	AAACATCTCC	CCTCTCACCT	1080
CATTCCCCCA	CCTATCCCCA	TTCTCTGCCT	GTACTGAAGC	TGAAATGCAG	GAAGTGGTGG	1140
CAAAGGTTTA	TTCCAGAGAA	GCCAGGAAGC	CGGTCATCAC	CCAGCCTCTG	AGAGCAGTTA	1200
CTGGGGTCAC	CCAACCTGAC	TTCTCTGCC	ACTCCCCGCT	GTGTGACTTT	GGGCAAGCCA	1260
AGTGCCCTCT	CTGAACCTCA	GTTCCTCAT	CTGCAAAATG	GGAACAATGA	CGTGCCTACC	1320
TCTTAGACAT	GTGTGAGGA	GACTATGATA	TAACATGTGT	ATGTAAATCT	TCATGTGATT	1380
GTCATGTAAG	GCTTAACACA	GTGGGTGGTG	AGTTCTGACT	AAAGGTTACC	TGTTGTCGTG	1440
AAAAAAAAAA	AAAA					1454

(2) INFORMATION FOR SEQ ID NO:3:

- (i) SEQUENCE CHARACTERISTICS:
 - (A) LENGTH: 19 base pairs
 - (B) TYPE: nucleic acid
 - (C) STRANDEDNESS: single
 - (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:3:

CCGCAGATT AGGTGACAC

(2) INFORMATION FOR SEQ ID NO:4:

- (i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 19 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:4:

GGCCTCTAAC ACGACTCAC

19

(2) INFORMATION FOR SEQ ID NO:5:

- (i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 17 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:5:

TACCACTACA ATGGATG

17

(2) INFORMATION FOR SEQ ID NO:6:

- (i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 20 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:6:

GTTGAAGTGA ACTTGCGGGC

20

(2) INFORMATION FOR SEQ ID NO:7:

- (i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 281 amino acids
(B) TYPE: amino acid
(C) STRANDEDNESS: not relevant
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:7:

Met Ser Ala Leu Leu Ile Leu Ala Leu Val Gly Ala Ala Val Ala Phe
1 5 10 15

Pro Val Asp Asp Asp Lys Ile Val Gly Gly Tyr Met Met Thr Arg
20 25 30

Tyr Ala Arg Thr Cys Arg Glu Ser Ser Val Pro Tyr Gln Val Ser Leu
35 40 45

Asn Ala Gly Tyr His Phe Cys Gly Gly Ser Leu Ile Asn Asp Gln Trp
50 55 60

Val Val Ser Ala Ala His Cys Tyr Lys Tyr Arg Ile Gln Val Arg Leu
65 70 75 80

Gly Glu His Asn Met Met Thr Arg Tyr Ala Arg Ile Asn Val Leu Glu
85 90 95

Gly Asn Glu Gln Phe Val Asp Ser Ala Lys Ile Ile Arg His Pro Asn
100 105 110

Tyr Asn Ser Trp Thr Leu Asp Asn Asp Ile Met Leu Ile Lys Leu Ala
115 120 125

Ser Pro Val Thr Leu Met Met Thr Arg Tyr Ala Arg Asn Ala Arg Val
130 135 140

Ala Ser Val Pro Leu Pro Ser Ser Cys Ala Pro Ala Gly Thr Gln Cys
145 150 155 160

Leu Ile Ser Gly Trp Gly Asn Thr Leu Ser Asn Gly Val Asn Asn Pro
165 170 175

Asp Leu Leu Gln Cys Val Asp Ala Pro Val Leu Pro Gln Ala Met Met
180 185 190

Thr Arg Tyr Ala Arg Asp Cys Glu Ala Ser Tyr Pro Gly Asp Ile Thr
195 200 205

Asn Asn Met Ile Cys Val Gly Phe Leu Glu Gly Gly Lys Asp Ser Cys
210 215 220

Gln Gly Asp Ser Gly Gly Pro Val Val Cys Asn Gly Glu Leu Gln Gly
225 230 235 240

Ile Val Ser Trp Gly Tyr Met Met Thr Arg Tyr Ala Arg Gly Cys Ala
245 250 255

Gln Pro Asp Ala Pro Gly Val Tyr Thr Lys Val Cys Asn Tyr Val Asp
260 265 270

Trp Ile Gln Asn Thr Ile Ala Asp Asn
275 280

(2) INFORMATION FOR SEQ ID NO:8:

- (i) SEQUENCE CHARACTERISTICS:
- (A) LENGTH: 299 amino acids
 - (B) TYPE: amino acid
 - (C) STRANDEDNESS: not relevant
 - (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:8:

Leu Glu Leu His Pro Leu Leu Gly Gly Arg Thr Trp Arg Ala Ala Arg
1 5 10 15

Asp Ala Asp Gly Cys Glu Ala Leu Gly Thr Val Ala Val Pro Phe Asp
20 25 30

Asp Asp Asp Lys Ile Val Gly Gly Tyr His Ser Thr Arg Tyr Ile Val
35 40 45

Asx Thr Cys Glu Asn Ser Leu Pro Tyr Gln Val Ser Leu Asn Ser Gly
50 55 60

Ser His Phe Cys Gly Gly Ser Leu Ile Ser Glu Gln Trp Val Val Ser
65 70 75 80

Ala Ala His Cys Tyr Lys Thr Arg Ile Gln Val Arg Leu Gly Glu His
85 90 95

Asn His Ser Thr Arg Tyr Ile Val Asx Ile Lys Val Leu Glu Gly Asn
100 105 110

Glu Gln Phe Ile Asn Ala Ala Lys Ile Ile Arg His Pro Lys Tyr Asn
115 120 125

Arg Asp Thr Leu Asp Asn Asp Ile Met Leu Ile Lys Leu Ser Ser Pro
130 135 140

Ala Val Ile His Ser Thr Arg Tyr Ile Val Asx Asn Ala Arg Val Ser
145 150 155 160

Thr Ile Ser Leu Pro Thr Ala Pro Pro Ala Ala Gly Thr Glu Cys Leu
165 170 175

Ile Ser Gly Trp Gly Asn Thr Leu Ser Phe Gly Ala Asp Tyr Pro Asp
180 185 190

Glu Leu Lys Cys Leu Asp Ala Pro Val Leu Thr Gln Ala His Ser Thr
195 200 205

Arg Tyr Ile Val Asx Glu Cys Lys Ala Ser Tyr Pro Gly Lys Ile Thr
210 215 220

Asn Ser Met Phe Cys Val Gly Phe Leu Glu Gly Gly Lys Asp Ser Cys
225 230 235 240

Gln Arg Asp Ser Gly Gly Pro Val Val Cys Asn Gly Gln Leu Gln Gly
245 250 255

Val Val Ser Trp Gly His His Ser Thr Arg Tyr Ile Val Asx Gly Cys
260 265 270

Ala Trp Lys Asn Arg Pro Gly Val Tyr Thr Lys Val Tyr Asn Tyr Val
275 280 285

Asp Trp Ile Lys Asp Thr Ile Ala Ala Asn Ser
290 295

(2) INFORMATION FOR SEQ ID NO:9:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 286 amino acids
(B) TYPE: amino acid
(C) STRANDEDNESS: not relevant
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:9:

Met Lys Ile Cys Ile Phe Phe Thr Leu Leu Gly Thr Val Ala Ala Phe
1 5 10 15

Pro Thr Glu Asp Asn Asp Asp Arg Ile Val Gly Gly Tyr Arg Asn Thr
20 25 30

Arg Tyr Pro Val Asx Thr Cys Gln Glu His Ser Val Pro Tyr Gln Val
35 40 45

Ser Leu Asn Ala Gly Ser His Ile Cys Gly Ser Leu Ile Thr Asp
50 55 60

Gln Trp Val Leu Ser Ala Ala His Cys Tyr His Pro Gln Leu Gln Val
65 70 75 80

Arg Leu Gly Glu His Asn Arg Asn Thr Arg Tyr Pro Val Asx Ile Tyr
85 90 95

Glu Ile Glu Gly Ala Glu Gln Phe Ile Asp Ala Ala Lys Met Ile Leu
100 105 110

His Pro Asp Tyr Asp Lys Trp Thr Val Asp Asn Asp Ile Met Leu Ile
115 120 125

Lys Leu Lys Ser Pro Ala Thr Leu Arg Asn Thr Arg Tyr Pro Val Asx
130 135 140

Asn Ser Lys Val Ser Thr Ile Pro Leu Pro Gln Tyr Cys Pro Thr Ala
145 150 155 160

Gly Thr Glu Cys Leu Val Ser Gly Trp Gly Val Leu Lys Phe Gly Phe
165 170 175

Glu Ser Pro Ser Val Leu Gln Cys Leu Asp Ala Pro Val Leu Ser Asp
180 185 190

Ser Arg Asn Thr Arg Tyr Pro Val Asx Val Cys His Lys Ala Tyr Pro
195 200 205

Arg Gln Ile Thr Asn Asn Met Phe Cys Leu Gly Phe Leu Glu Gly Gly
210 215 220

Lys Asp Ser Cys Gln Tyr Asp Ser Gly Gly Pro Val Val Cys Asn Gly
225 230 235 240

Glu Val Gln Gly Ile Val Ser Trp Gly Asp Arg Asn Thr Arg Tyr Pro
245 250 255

Val Asx Gly Cys Ala Leu Glu Gly Lys Pro Gly Val Tyr Thr Lys Val
260 265 270

Cys Asn Tyr Leu Asn Trp Ile Gln Gln Thr Val Ala Ala Asn
275 280 285

(2) INFORMATION FOR SEQ ID NO:10:

- (i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 271 amino acids
(B) TYPE: amino acid
(C) STRANDEDNESS: not relevant
(D) TOPOLOGY: linear
- (ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:10:

Ala Ala Phe Ala Thr Glu Asp Asp Lys Ile Val Gly Gly Tyr Ser Ser
1 5 10 15

Thr Arg Tyr Pro Ile Ile Glu Cys Lys Ala Tyr Ser Gln Pro His Gln
20 25 30

Val Ser Leu Asn Ser Gly Tyr His Phe Cys Gly Ser Leu Val Asn
35 40 45

Glu Asn Trp Val Val Ser Ala Ala His Cys Tyr Gln Ser Arg Val Glu
50 55 60

Val Arg Leu Gly Glu His Asn Ser Ser Thr Arg Tyr Pro Ile Ile Ile
65 70 75 80

Gln Val Thr Glu Gly Ser Glu Gln Phe Ile Ser Ser Ser Arg Val Ile
85 90 95

Arg His Pro Asn Tyr Ser Ser Tyr Asn Ile Asp Asn Asp Ile Met Leu
100 105 110

Ile Lys Leu Ser Lys Pro Ala Thr Leu Ser Ser Thr Arg Tyr Pro Ile
115 120 125

Ile Asn Thr Tyr Val Gln Pro Val Ala Leu Pro Thr Ser Cys Ala Pro
130 135 140

Ala Gly Thr Met Cys Thr Val Ser Gly Trp Gly Asn Thr Met Ser Ser
145 150 155 160

Thr Ala Asp Lys Asn Lys Leu Gln Cys Leu Asn Ile Pro Ile Leu Ser
165 170 175

Tyr Ser Ser Ser Thr Arg Tyr Pro Ile Ile Asp Cys Asn Asn Ser Tyr
180 185 190

Pro Gly Met Ile Thr Asn Ala Met Phe Cys Ala Gly Tyr Leu Glu Gly
195 200 205

Gly Lys Asp Ser Cys Gln Gly Asp Ser Gly Gly Pro Val Val Cys Asn
210 215 220

Gly Glu Leu Gln Gly Val Val Ser Trp Gly Tyr Ser Ser Thr Arg Tyr
225 230 235 240

Pro Ile Ile Gly Cys Ala Glu Pro Gly Asn Pro Gly Val Tyr Ala Lys
245 250 255

Val Cys Ile Phe Asn Asp Trp Leu Thr Ser Thr Met Ala Thr Tyr
260 265 270

(2) INFORMATION FOR SEQ ID NO:11:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 291 amino acids
- (B) TYPE: amino acid
- (C) STRANDEDNESS: not relevant
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:11:

Met Arg Ala Pro His Leu His Leu Ser Ala Ala Ser Gly Ala Arg Ala
1 5 10 15

Leu Ala Lys Leu Leu Pro Leu Leu Met Ala Gln Leu Trp Ala Ala Glu
20 25 30

Ala Ala Leu Leu Pro Gln Asn Asp Thr Arg Leu Asp Pro Glu Ala Tyr
35 40 45

Gly Ala Asn Glu Ser Pro Cys Ala Arg Gly Ser Gln Pro Trp Gln Val
50 55 60

Ser Leu Phe Asn Gly Leu Ser Phe His Cys Ala Gly Val Leu Val Asp
65 70 75 80

Gln Ser Trp Val Leu Thr Ala Ala His Cys Gly Asn Lys Pro Leu Trp
85 90 95

Ala Arg Val Gly Asp Asp His Asn Glu Ser Leu Leu Leu Gln Gly
100 105 110

Glu Gln Leu Arg Arg Thr Thr Arg Ser Val Val His Pro Lys Tyr His
115 120 125

Gln Gly Ser Gly Pro Ile Leu Pro Arg Arg Thr Asp Glu His Asp Leu
130 135 140

Met Leu Leu Lys Leu Ala Arg Pro Val Val Pro Asn Glu Ser Gly Pro
145 150 155 160

Arg Val Arg Ala Leu Gln Leu Pro Tyr Arg Cys Ala Gln Pro Gly Asp
165 170 175

Gln Cys Gln Val Ala Gly Trp Gly Thr Thr Ala Ala Arg Arg Val Lys
180 185 190

Tyr Asn Lys Gly Leu Thr Cys Ser Ser Ile Thr Ile Leu Ser Pro Lys
195 200 205

Asn Glu Ser Glu Cys Glu Val Phe Tyr Pro Gly Val Val Thr Asn Asn
210 215 220

Met Ile Cys Ala Gly Leu Asp Arg Gly Gln Asp Pro Cys Gln Ser Asp
225 230 235 240

Ser Gly Gly Pro Leu Val Cys Asp Glu Thr Leu Gln Gly Ile Leu Ser
245 250 255

Trp Gly Val Tyr Asn Glu Ser Pro Cys Gly Ser Ala Gln His Pro Ala
260 265 270

Val Tyr Thr Gln Ile Cys Lys Tyr Met Ser Trp Ile Asn Lys Val Ile
275 280 285

Arg Ser Asn
290